

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
4 November 2004 (04.11.2004)

PCT

(10) International Publication Number
WO 2004/094636 A1

(51) International Patent Classification⁷: C12N 15/11,
A01K 67/027, C12N 15/861, C12Q 1/68, C12N 15/10,
5/10

(21) International Application Number:
PCT/EP2003/004362

(22) International Filing Date: 24 April 2003 (24.04.2003)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicants (for all designated States except US): GALA-
PAGOS GENOMICS N.V. [BE/BE]; Generaal de
Wittelaan 11 A3, B-2800 MECHELEN (BE). VAN DER
SCHUEREN, Jan [BE/BE]; Spaarstraat 56, B-3010
KESSEL-LO (BE).

(72) Inventors; and

(75) Inventors/Applicants (for US only): ARTS, Gerardus,
Johannes, Franciscus [NL/NL]; Wikkestraat 46, NL-2403
ES ALPHEN AAN DE RIJN (NL). LAMBRECHT,
Mark, Jacques, Yvonne [BE/BE]; Bukenstraat 89,
B-1910 BUKEN (BE). DJOKIC, Kristina [NL/NL];
Joseph Haydnlaan 58, NL-2324 AT LEIDEN (NL).

CLASEN, Remko, Johannes [NL/NL]; Middelweg 15
H, NL-2312 KE LEIDEN (NL). MESIC, Emir [NL/NL];
Koningstraat 1, NL-2316 CA LEIDEN (NL). GRIF-
FIOEN, Sandra [NL/NL]; Maresingel 47A, NL-2316 HD
LEIDEN (NL). BERGS, Carolina, Johanna, Leonarda
[NL/NL]; Van Vollenhovenplein 102, NL-2313 EE LEI-
DEN (NL).

(74) Agent: VERNOUT, Robert; Arnold & Siedsma, Sweel-
inckplein 1, NL-2517 GK THE HAGUE (NL).

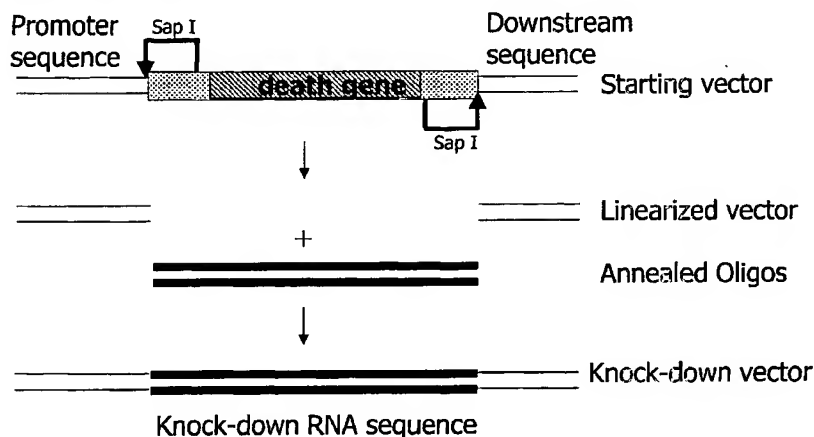
(81) Designated States (national): AE, AG, AL, AM, AT, AU,
AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU,
CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH,
GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC,
LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW,
MX, MZ, NI, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD,
SE, SG, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US,
UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States (regional): ARIPO patent (GH, GM,
KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),
Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM),
European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO,

[Continued on next page]

(54) Title: EFFECTIVE SIRNA KNOCK-DOWN CONSTRUCTS

Vector development



(57) Abstract: The present invention provides polynucleotides comprising an RNA sequence comprising a first stretch of 21 consecutive nucleotides and a second stretch of 21 consecutive nucleotides, complementary to the first stretch of 21 consecutive nucleotides, wherein in first stretch of 21 consecutive nucleotides: in the 5' -> 3' direction, the first nucleotide is an A-nucleotide, the second nucleotide is a C-nucleotide and the last nucleotide is a C-nucleotide; and no stretches of four or more consecutive identical nucleotides are present; the total number of G- and C-nucleotides is between 33-71% of the total number of nucleotides and in the 5' -> 3' direction, consecutive nucleotides 3-21 are homologous to a RNA-molecule. Vectors encoding the polynucleotide of the present invention are also provided. Furthermore libraries of the polynucleotides and vectors encoding and methods to produce vectors and libraries are given.

WO 2004/094636 A1



SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *with international search report*

Effective siRNA knock-down constructs

The present invention relates to polynucleotide constructs, methods for their preparation, and preparations
5 for their use in methods that lower the amount of RNA and/or protein production in cells based on the intracellular expression of small interfering polyribonucleic acid molecules.

Pharmaceutical companies are interested in reliable
10 knockdown based technologies since their drug screens with small molecules are based on inhibiting the activity and effect of an expressed protein. Therefore, blocking expression or function of a potential target, either through screening in a cellular assay or through single gene
15 validation will provide an important data set regarding drug-ability of the target early on in the drug development process. This data set forms a strong basis for the start of a drug development program, based on a compound, antibody or biological, with the aim to develop an effective therapy.

20 Various knockdown or knockout approaches are used to study gene function in mammalian cells (e.g. antisense, antibodies, ribozymes, aptamers, zinc finger proteins, chimeric RNA-DNA oligos, etc.). However, these technologies are not robust and efficient nor they can be generically
25 applied to all genes and all cell types.

RNA interference (RNAi) is the post-transcriptional process of gene silencing mediated by double stranded RNA (dsRNA) that is homologous in sequence to the silenced RNA and is observed in animals and plants. The dsRNA is processed
30 into 21-23 nucleotides (nts) molecules, called small interfering RNAs (siRNAs), which guide the sequence-specific degradation of the target RNA.

RNAi provides researchers with an additional genetic tool to study gene functions. In C. elegans, chromosomes I and III have now systematically been analyzed for phenotypic effects. The RNAi approach creates extra possibilities in developmental studies. Classical knockouts with lethal effects during development could never be analyzed in later developmental stages. With RNAi, the onset of the effect may be varied and roles in later stages of development may be studied.

The use of RNAi in mammalian cells has been problematic since introduction of long (>30 base pairs) dsRNA results in two major intracellular responses: activation of the double stranded RNA dependent protein kinase PKR, which results in a general block of protein synthesis. Long dsRNA is also known to activate the interferon-induced (2'-5') oligoadenylate synthetase. Upon activation, this enzyme polymerizes ATP into 2'-5'-linked nucleotide oligomers (also indicated by 2-5A). The 2-5A oligomers activate the ribonuclease RNase L that results in RNA degradation.

Recently, it has been demonstrated that RNAi can be used in a panel of mammalian cell lines. The approach is based on direct transfection of the 21-23 nts siRNA duplexes into the cells. This circumvents the intracellular responses mentioned above and results in sequence-specific silencing of endogenous and heterologous genes.

An important bottleneck in the siRNA transfection approach is its limited applicability to target different cell types, especially primary cells. Primary cells are closest to the in vivo situation and often have the highest physiological relevance. Non-viral DNA or siRNA transfection technologies have severe limitations with regard to these cells and are not efficient and reliable. Practical use of these approaches needs significant optimisation of

conditions, and in general lack the robustness necessary for large-scale applications. The gene transfer reagents used are often toxic, yielding lower levels of viable transduced cells. In essence, they do not allow a generic siRNA application for a wide variety of cell types, including primary cell types such as T cells, B cells, mast cells, endothelial cells, synoviocytes and lung epithelial cells. Furthermore, transfection of the siRNA gives a short knock-down effect. For a prolonged knock-down effect in cells several additional transfections are necessary.

Viral constructs encoding the siRNA molecule circumvent the problems described above. WO03020391 describes adenoviral vectors which express hairpin RNAs that are further processed to siRNAs. Infection of cell lines, or primary human cells, with these viruses leads to an efficient, sequence-specific, and prolonged reduction of the corresponding target mRNA, resulting in a functional knock-down of the encoded protein.

However, not every siRNA can effectively downregulate a gene. siRNAs directed to different regions of a mRNA can result in different levels of gene silencing. It is estimated that between 25% and 75% of siRNAs are effective. Some RNAi's have a 10-fold effect, some have seen a 50-fold effect, and some don't work at all. For example, a siRNA directed against vimentin (nt 346-368 from Genbank (NCBI): NM_003380 relative to start codon) did not give a silencing effect. Three other siRNAs against vimentin were designed and these all gave an effective gene silencing effect (nt 1145-1167, nt 863-885 and nt 1037-1059 from NM_003380 relative to start codon, (Elbashir et al 2001)). A procedure for designing siRNAs for efficiently inducing RNAi in mammalian cells has been suggested (Elbashir et al 2002). However, a siRNA against c-myc, designed according to this protocol was ineffective in silencing. and did not show a reduction in

protein expression Jarvis and Ford (2002). This clearly demonstrates that target site selection is critical for the effective induction of RNAi by siRNAs.

The present invention provides polynucleotides that are
5 very effective in the silencing a RNA molecule.

The present invention provides polynucleotide comprising an RNA sequence comprising a first stretch of 21 consecutive nucleotides and a second stretch of 21 consecutive nucleotides, complementary to the first stretch of 21
10 consecutive nucleotides, wherein in first stretch of 21 consecutive nucleotides:

- a. in the 5' -> 3' direction, the first nucleotide is an A-nucleotide , the second nucleotide is a C-nucleotide and the last nucleotide is a C-nucleotide; and
- 15 b. no stretches of four or more consecutive identical nucleotides are present;
- c. the total number of G- and C-nucleotides is between 33-71% of the total number of nucleotides;
- d. in the 5' -> 3' direction, consecutive nucleotides 3-
20 21 are homologous to a RNA-molecule.

The total number of G- and C-nucleotides can be 33, 35, 40, 45, 50, 55, 60, 65, 70, and 71 % of the total number of nucleotides. The consecutive nucleotides 3-21 in the 5' -> 3' direction of the first stretch of 21 consecutive nucleotides
25 are homologous to the RNA-molecule to be silenced.

"Homologous" to the RNA-molecule can mean that nucleotides 3-21 of the stretch of 21 consecutive nucleotide are identical to a stretch of 19 consecutive nucleotides present in the RNA-molecule to be silenced, or that 95%, 90%,
30 85% or 80% of the nucleotides 3-21 of the stretch of 21 consecutive nucleotide are identical to a stretch of 19 consecutive nucleotides present in the RNA-molecule to be silenced. A polynucleotide with all 21 nucleotides of the

stretch of 21 consecutive nucleotide identical to a stretch of 21 consecutive nucleotides present in the RNA-molecule to be silenced is also possible. In a suitable manner of the present invention the first and second stretch of 21 consecutive nucleotides form a double stranded RNA molecule.

In a preferred embodiment the invention provides polynucleotides wherein the RNA-molecule is a human RNA molecule. The RNA molecule will be silenced by the polynucleotides of the present invention. The RNA molecule to be silenced can be a drugable gene. Drugable genes are genes with a pharmaceutical value, ie that can be used to discover and develop small molecule drugs. New drugable genes can be found on the bases on their similarity to proteins which have proven amenable to small molecule compound development in the past. Known drugable genes include but are not limited to: G-protein coupled receptors (GPCRs), ion channels, nuclear hormone receptors, kinases, phosphatases, proteases and other enzymes.

In a preferred embodiment the invention provides polynucleotides wherein in the first stretch of 21 consecutive nucleotides no stretches of three or more consecutive A- nucleotides are present.

In another preferred embodiment the invention provides polynucleotides wherein in the first stretch of 21 consecutive nucleotides no stretches of three or more consecutive U-nucleotides are present

In a preferred embodiment the invention provides polynucleotides, wherein, in the 5' -> 3' direction, consecutive nucleotides 3-21 of the first stretch of 21 consecutive nucleotides are unique. These unique oligonucleotides are homologous to a stretch of 21 consecutive nucleotides that is found only once in known sequences. Sequence databases can be searched to check if

these sequences occur only once. Known databases include the EST database, the EMBL nucleotide sequence database, GenBank, and the Entrez nucleotide database, but there are many more sequence databases and all of these can be used. The

5 polynucleotide with the unique consecutive nucleotides 3-21 in the 5' -> 3' direction, of the first stretch of 21 consecutive nucleotides will silence only one specific mRNA molecule. Alternatively, the consecutive nucleotides 3-21 in the 5' -> 3' direction, of the first stretch of 21

10 consecutive nucleotides are found more than once in sequence databases. These polynucleotides can silence alternative gene transcripts of the same sequence or alternative splicing variants. In addition, the stretch of 21 consecutive nucleotides can be designed in such a way that the sequence

15 is found in more than one member of a family of proteins. The polynucleotides containing this stretch of 21 consecutive nucleotides can silence more than one member of a family of proteins.

According to another preferred embodiment of the

20 invention polynucleotides are provided wherein, in the 5' -> 3' direction, consecutive nucleotides 3-21 of the first stretch of 21 consecutive nucleotides are homologous to a sequence positioned at least 75 nucleotides downstream of the translation initiation site of the transcribed RNA molecule

25 encoding a polypeptide.

According to a preferred embodiment of the invention polynucleotides are provided wherein, in the 5' -> 3' direction, consecutive nucleotides 3-21 the first stretch of 21 consecutive nucleotides homologous are to a sequence

30 positioned at least upstream of the translation termination site of the transcribed RNA molecule encoding a polypeptide.

According to a preferred embodiment of the invention polynucleotides are provided, wherein the RNA sequence also

comprises a linker sequence linking the first stretch of 21 consecutive nucleotides with the second stretch of 21 consecutive nucleotides.

Preferably the linker sequence is 4-30 nucleotides long, more preferably 5-15 nucleotides long and most preferably 8 nucleotides long.

In a most preferred embodiment the linker sequence is UUGCUAUA (SEQ ID NO: 1).

According to a preferred embodiment the invention provides polynucleotides, wherein the first stretch of 21 consecutive nucleotides is selected from a group consisting of SEQ ID NO: 2-11888.

According to another preferred embodiment the invention provides polynucleotides, wherein the first stretch of 21 consecutive nucleotides is selected from a group consisting of SEQ ID NO: 1342, 1338, 1343, 633, 635.

Another embodiment of the invention provides vector capable of transfecting a host cell and comprising a sequence encoding the polynucleotides of the present invention and a promoter sequence operatively linked to the sequence encoding the polynucleotide.

Preferably the promoter is a microRNA promoter, more preferably a let-7 promoter.

More preferably the promoter is a promoter recognized by RNA Polymerase III, more preferably U6 small nuclear RNA. A person skilled in the art can use other promoters recognized by Polymerase III for the vectors of the present invention such as, H1, tRNA, snRNA, VA RNA, 5S rRNA.

According to a preferred embodiment of the invention vectors are provided, wherein the vector is an adenoviral vector, preferably the adenoviral vector is replication defective. The replication defective adenoviral vectors are E1-deleted, and/or E1 and E2A deleted. The adenoviral vectors

include the E1-deleted adenoviral serotype 5 vectors. Vectors may also be prepared from other adenoviral serotypes and corresponding packaging cells that include sequences for viral proteins deleted from such vector backbones. A suitable approach of the present invention has an adenoviral vector/packaging cell wherein the packaging cell and vector do not include any overlapping adenoviral sequences, which overlap would provide the statistical possibility of the production of replication competent adenoviral particles.

Packaging cells useful in the production of such vectors include the 293 and 911 cells, with the most suitable cells being the PER.C6 cell line. The modified PER.C6/E2A cell line is especially suitable. It complements the E1, E2A deleted adenoviral vector constructs, with non-overlapping adenoviral E1, E2A sequences. Other viral vector systems can be used such as the retroviral vector systems. Retroviruses are integrating viruses that infect dividing cells, and their construction is known in the art. Retroviral vectors can be constructed from different types of retrovirus, such as, MoMuLV ("murine Moloney leukemia virus" MSV ("murine Moloney sarcoma virus"), HaSV ("Harvey sarcoma virus"); SNV ("spleen necrosis virus"); RSV ("Rous sarcoma virus") and Friend virus. Lentivirus vector systems such as human immunodeficiency virus (HIV) or equine lentivirus may also be used in the practice of the present invention. Another suitable viral vector system is the adeno-associated virus ("AAV"). The AAV viruses are DNA viruses of relatively small size that integrate, in a stable and site-specific manner, into the genome of the infected cells. They are able to infect a wide spectrum of cells without inducing any effects on cellular growth, morphology or differentiation, and they do not appear to be involved in human pathologies. It is also possible to introduce a DNA vector in vivo as a naked DNA

plasmid. Naked DNA vectors for gene therapy can be introduced into the desired host cells by methods known in the art, e.g., transfection, electroporation, microinjection, transduction, cell fusion, DEAE dextran, calcium phosphate
5 precipitation, use of a gene gun, or use of a DNA vector transporter.

The invention further provides libraries of polynucleotide of the present invention.

Preferably the library of the present invention provides
10 vectors according to the present invention. The vectors may comprise plasmids, naked RNA. Preferably, the vectors are viral vectors preferably selected from a group consisting of AAV, Lentivirus or Retrovirus. Alternatively, more than one vector thereby introducing more than one stretch of 21
15 consecutive nucleotides can be introduced into a single host cell.

More preferred vectors are adenoviral vectors, preferably the adenoviral vectors are replication defective. Replication deficient vectors may be multiplied in a
20 packaging cell having complementary sequences to the sequence contained in the vector itself.

The present invention also provides means to prepare libraries of polynucleotides and vectors as described herein. These libraries may be prepared as single element,
25 compartmentalized, or discrete elements. Alternatively, a library comprising pools of vectors may be prepared. Methods are described for making a vector comprising the synthesis a forward primer and a reverse primer.

- a. The forward primer is synthesized with the following
30 sequence: in the 5' -> 3' direction:
- i. the nucleotides ACC
 - ii. a DNA sequence corresponding to a sequence selected from the group consisting of SEQ ID NO: 2-11888

- iii. the nucleotides TTGCTATA
 - iv. the antisense sequence corresponding to the DNA sequence from step a-ii
 - v. the nucleotides TTT
- 5 b. the reverse primer is synthesized with the following sequence in the 5' -> 3' direction
- i. the nucleotides TAAAAA
 - ii. a DNA sequence corresponding to a sequence selected from the group consisting of SEQ ID NO: 2-11888
- 10 iii. the nucleotides TATAGCAA
- iv. the antisense sequence corresponding to the DNA sequence from step b-ii
- c. The primers are annealed and cloned in plasmid pKD122, thereby exchanging the ccdB sequences for the annealed
- 15 primers.

The orientation of the sequence encoding the sequence selected from the group consisting of SEQ ID NO: 2-11888 can be reversed. The forward and reverse primers look then as follows:

- 20 a. forward primer with the following sequence in the 5' -> 3' direction:
- i. the nucleotides ACC
 - ii. a DNA sequence corresponding to a sequence selected from the group consisting of SEQ ID NO: 2-11888
- 25 iii. the nucleotides TTGCTATA
- iv. the antisense sequence corresponding to the DNA sequence from step a-ii
 - v. the nucleotides TTT
- b. a reverse primer with the following sequence in the 5' -> 3' direction
- 30 i. the nucleotides TAAAAA
- ii. a DNA sequence corresponding to a sequence selected from the group consisting of SEQ ID NO: 2-11888

iii. the nucleotides TATAGCAA

iv. the antisense sequence corresponding to the DNA sequence from step b-ii

The invention further provides methods of determining
5 the function of a naturally occurring polynucleotide sequence comprising transfecting a host cell with a vector according to claims 12-15, the vector transcribing a polynucleotide sequence according to claims 1-11 and detecting a change in cellular phenotype.

10 According to a preferred embodiment of the present inventions methods of determining the function of a naturally occurring polynucleotide sequence in a high throughput setting are provided wherein,

- 15 a. providing a library of vectors according to claims 16 - 19
- b. transducing a host cell with the vectors of step (a),
- c. expressing in the host cell the product(s) of the vectors of step (a),
- d. thereby altering a phenotype of the host,
- 20 e. identifying the altered phenotype and,
- f. assigning a function to the naturally occurring polynucleotide sequence (s).

The libraries according to the invention may be used to assist in the elucidation of the functions of host cell RNA
25 molecules including the polynucleotide of the present invention residing in each compartment of said library. In other words, determining the function of a naturally occurring polynucleotide sequence comprising transfecting a host cell with a vector according to the invention, the
30 vector, encoding a RNA molecule including a stretch of 21 consecutive nucleotides homologous to a portion of the naturally occurring polynucleotide and detecting a change in cellular phenotype. Each vector in the library may be

introduced into one or more cells and changes in protein expression, or phenotype observed. Methods are described for infecting a host with the adenoviral vectors that express the RNA molecules including the stretch of 21 consecutive
5 nucleotides in the host, identifying an altered phenotype induced in the host by the knockdown of the expressed RNA molecules, and thereby assigning a function to the product(s) encoded by the expressed RNA molecules. The methods can be fully automated and performed in a multiwell format to allow
10 for convenient high throughput analysis of expressed RNA molecules.

The term "homologous" to a RNA molecule refers to a nucleic acid having a nucleotide sequence of which 95%, 90%, 85% or 80% of the nucleotides 3-21 of the stretch of 21
15 consecutive nucleotide are identical to a stretch of 19 consecutive nucleotides present in the RNA-molecule to be silenced. For example, with 95% homology the nucleotide sequence is identical to reference nucleotide, i.e. the RNA molecule, except that the nucleotide sequence may include up
20 to five point mutations per each 100 nucleotides of the reference nucleotide sequence. In other words, to obtain a nucleic acid having a nucleotide sequence of at least 95% homology to a reference nucleotide sequence, up to 5% of the nucleotides in the reference sequence may be deleted or
25 substituted with another nucleotide, or a number of nucleotides up to 5% of the total nucleotides in the reference sequence may be inserted into the reference sequence.

The term identical refers to two stretches of
30 nucleotides that have at each position of the stretch the same nucleotide, meaning that if one stretch has a A-nucleotide on position three, the other stretch of

nucleotides also has a A-nucleotide on position three. This is applicable for each nucleotide in both stretches.

The term polynucleotide refers to a nucleic acid sequence. A polynucleotide can be a DNA-, RNA-, peptide
5 nucleic acid sequence. It may have natural occurring nucleotides but also chemically modified nucleotides.

Brief Description of the Drawings

Figure 1 is a schematic representation of the cloning
10 strategy for vector construction showing utilization of Sapi sites and an E. coli death gene.

Figure 2 is a schematic representation of the vector construction for cloning. Adenoviral vector development for 56 nt inserts. The schematic presentation of the oligos used
15 for the vector construction is also given.

Figure 3 is a schematic presentation of the knock-down vector pKD122.

Figure 4 is a graph showing the functional knock-down of GNAS. Adenoviral constructs encoding hairpin siRNA targeted
20 against GNAS give a specific knock-down of GNAS on the functional level.

Figure 5 is a graph showing the functional knock-down of SYK as measured by the released beta-hexosaminidase.

Figure 6 shows a graph demonstrating that adenoviral
25 mediated knock-down of IKK β leads to reduced levels of phosphorylated I κ B α , an I κ B kinase substrate, upon TNF α induction. Western detection of IKK β and β -tubulin in the samples derived from U2OS cells (no virus infection, lanes 1-3; infected with adenovirus Ad-IKK- β , lanes 4-6) harvested at
30 different time points after 50 ng/ml TNF α stimulation (0 minutes, lanes 1, 4, or 10 minutes, lanes 2, 5, or 30 minutes, lanes 3, 6). Detection of phosphorylated I κ B α by

Bio-Plex phosphoprotein analysis of the same samples as described. The phosphorylation state of I κ B α is given as Mean Fluorescent Intensity (M.F.I.) at the vertical axis.

Figure 7: shows a graph demonstrating the functional
5 knock-down of MMP2. The effect of adenoviral knock-down was determined by RT-PCR and at the level of protein activity. Gelatin-degrading activity of MMP2 was measured using 10% gelatin-zymogram gels. For reference, an adenoviral cDNA expression construct encoding the MMP2 protein (Ad-MMP2) was
10 used and demonstrated the identity of the MMP2 band. The inverted image of the zymogram is presented.

The various aspects of the present invention are further illustrated in the following non-limiting Examples.

15

Example 1: selecting target sequences:

The polynucleotides of the present invention are built from the stretches of 21 consecutive nucleotides as explained in example 2. The design for the 21 consecutive nucleotide
20 sequences that form part of the polynucleotides of the present invention is fully automated. A program has been developed that looks for candidate sequences of 21 consecutive nucleotides. These 21 nucleotide sequences are searched in a target cDNA sequence. The program searches 21
25 nucleotide sequences according to the following scheme:

1. The search for the 21 nucleotide sequences starts 75
nucleotides downstream from the startcodon of the cDNA.
If a startcodon is not known or available, the search
30 starts 75 nt downstream of the first ATG. If no ATG is present, the 21 nucleotide sequences are searched starting from the first nucleotide of the cDNA sequence.

2. The cDNA sequence is scanned from 5' to 3' direction and when a 21 nucleotide sequence is found that conforms to a NM(N¹⁸)C pattern, it is retained (N=any nucleotide, M=A,C). This procedure is stopped 21 nucleotides before the end of the cDNA sequence.
3. At this point, the 21 nucleotide sequences that are found are attributed a score of 1000.
4. Then, the candidate 21 nucleotide sequence is checked for the presence of 4 consecutive identical nucleotides. If this pattern is found, this candidate 21 nucleotide sequence is rejected and the program starts at step 2 again.
5. If the 21 nucleotide sequence is not rejected in step 4, the percentage of G and C nucleotides of the candidate 21 nucleotide sequence is determined. The percentage of G and C nucleotides should be between 30 and 70%. If it is not between 30 and 70%, the 21 nucleotide sequence is rejected and the procedure starts at step 2 again. If the percentage is between 30 and 70%, the score of the candidate 21 nucleotide sequence is decreased according to the following rule :
- $$[(G - 0.5) * 100]$$
- with G = the fraction of G and C nucleotides of the 21 nucleotide sequence . The absolute value is subtracted from the score of the 21 nucleotide sequence.
6. The candidate 21 nucleotide sequence is checked for the presence of 3 consecutive A- or T-nucleotides. If this pattern is found, the score of the 21 nucleotide sequence is decreased with 5 units.
7. The 21 nucleotide sequence is then searched against a number of public sequence databases. The blast algorithm, (the National Center for Biotechnology Information (NCBI)) is used to match the 21 nucleotide sequence against ;

- a. -a human cDNA database, consisting of the complete RefSeq collection. For every extra cDNA sequence, apart from the target cDNA sequence, that matches the 21 nucleotide sequence, the score of the 21 nucleotide sequence is decreased. The score is decreased with 10 units for the first additional matching cDNA sequence and with another 4 units for every additional matching cDNA sequence.
- b. -the mouse RefSeq database
- 10 c. -the rat RefSeq database The score of the 21 nucleotide sequence is increased with 35 units for every matching mouse or rat cDNA sequence.
8. The 21 nucleotide sequence is then searched against the human 'Expressed Sequence Tag' (EST) database. If the 21 nucleotide sequence does not match any ESTs, its score is decreased by 10 units. If the 21 nucleotide sequence matches perfectly to an EST, the complete EST is matched against the target cDNA sequence. The program used for this pairwise sequence comparison is blast2seq, available from NCBI. From this pairwise sequence comparison, one can determine the sequence similarity of the EST and cDNA (p-value). If the p-value is e^{-100} or lower than the EST is similar to the cDNA target sequence, if the p-value is more than e^{-100} than the EST is dissimilar to the cDNA sequence. If the 21 nucleotide sequence does match ESTs, its score is decreased by 50 units for the first dissimilar EST and another 5 units for every extra dissimilar EST. The score of the 21 nucleotide sequence is increased with one unit for every EST sequence similar to the target cDNA sequence.
9. Procedure 2 until 8 is repeated until the end of the cDNA is reached or when 25 oligonucleotide sequences homologous to the target cDNA have been identified when the cDNA is

shorter than 2000 basepairs, or when the cDNA sequences is longer than 2000 basepairs, 50 nucleotide sequence are searched, and if the cDNA sequence is longer than 3000 basepairs, 60 nucleotide sequence are searched.

- 5 10. Each 21 nucleotide sequence has a score. All possible combinations of 4 of the designed 21 nucleotide sequence (so called 'sets') are made. For each set a combined score is calculated according to the following formula :

$$F(\text{oligo } 1\text{-}n) = T + ((\Delta_1 * \dots * \Delta_{n-1}) * W / S^{n-1}) \text{ with}$$

- 10 $T =$ sum of scores of all 21 nucleotide sequence in set
 $\Delta_n =$ difference of position on target cDNA of oligo $n+1$ and position on target cDNA of oligo n in set. Oligo 1 to n are sorted according to their position on the target cDNA sequence

- 15 $W =$ weighing factor to weigh importance of second term in above formula, W is set to 10000.

$S =$ scaling factor to normalize for size of cDNA (= size of open reading frame on target cDNA sequence: (position stopcodon - position startcodon)).

- 20 When sets are made of 5 or another number of 21 nucleotide sequence then the formula above can still be used to calculate the combined score.

11. For all computed sets and their attached function outcomes, the set with the highest function outcome is
25 kept and marked in the output report as the 'best' set. This 'best' set of four is suggested to the end-user on the generated output report (see point 12)

12. From this best set of 4 21 nucleotide sequences, obtained in step 10, the 3 highest scoring 21 nucleotide
30 sequences are used to design the oligos for the plasmid construction as described in example 2. The first and second nucleotide in the 21 nucleotide sequence are

exchanged for AC. These 21 nucleotide sequence are listed in Table 1.

13. Finally, an output report is generated. In this report, the cDNA sequence is graphically represented, together with all the designed 21 nucleotides sequence, and with the similar and dissimilar EST sequences. In a second part of the report, the designed 21 nucleotide sequences are listed with the characteristics mentioned in point 1 to 7. These include the matching human, mouse and rat transcript sequences, the number of ESTs, and the outcome of each of the pairwise comparisons to the target cDNA sequence.

The program performing the rules described above has been written in perl (practical extracting and reporting language). In our case extensive use has been made of the 'bioperl' bioinformatics toolkit (Perl, O'Reilly.com). However, a person skilled in the art can use another language to implement the rules set above in a program and use this program to design siRNA that give an effective knock-down effect.

20

Example 2: siRNA expression constructs:

Non viral siRNA expression constructs:

The construction of the siRNA expression constructs is depicted in figure 1. In short, oligos containing knock-down target sequences as depicted in table 1 (21 nt) are cloned in the knock-down vector, pKD122. This non-viral DNA expression plasmids can be used using DNA transfer methods known in the art, such as lipofectamine or PEI. The individual knockdown constructs for each gene can be pooled or can be used separately.

30

adenoviral siRNA expression constructs

The siRNA expression construct can also so be contained in viruses. The viruses can be made in an arrayed format, if

preferred. The arrayed viruses mediate expression of the siRNA constructs; each well contains a unique recombinant virus carrying a siRNA expression construct targeted against a gene, i.e. one target gene per well. Further details about
5 the concept of arrayed adenoviral vectors can be found in WO 9964582, US 6,340,595 and 6,413,776 (Arrayed adenoviral libraries for performing functional genomics).

In addition to the knock-down vector, pKD122 two other materials are needed for the generation of recombinant
10 adenovirus particles: a helper cosmid and a packaging cell line (see also WO9964582. US 6,340,595). The cosmid (pWE/Ad.AflIII-rITRΔE2A) contains the main part of the adenovirus serotype 5 genome (bp 3534-35953) from which the E2A gene is deleted. The Per.C6/E2A packaging cell line
15 (Crucell NV) is derived from human embryonic retina cells (HER) transfected with plasmids mediating the expression of the E1 and E2A genes. The adenoviral genes that are integrated into the genome of the PER.C6/E2A cell line share no homology with the adenoviral sequences on the knock-down
20 plasmid and the cosmid. Consequently, vector stocks that are free of replication competent adenoviruses (RCAs) are prepared.

To obtain viruses, the knockdown plasmid is co-transfected with the helper cosmid into a packaging cell line
25 PER.C6/E2A. Once these plasmids are transfected into the PER.C6/E2A cell line, the complete Ad5 genome is reconstituted by homologous recombination. The helper and knock-down plasmids contain homologous sequences (bp 3535-6093), which are a substrate for this recombination event.

30 **Design of oligos:**

Oligonucleotides are designed to be targeted against specific mRNAs. The selected target sequences are listed in

Table 1 and are used for the construction of knock-down adenoviral expression clones. Specific pairs of forward (F) and reverse (R) oligonucleotides are annealed together forming a duplexed structure that is used for cloning into the knock-down vector (see figure 2). The 56 nt oligos containing knock-down target sequences have the following structure :

the Forward oligonucleotide:

5'-ACC- N21*- TTGCTATA -N21--TTT-3'

Reverse oligonucleotide:

3'-N21-AACGATA- N21* -AAAAAT-5'

N21 are DNA sequences corresponding to the sequences as depicted in Table 1, N21* is the antisense sequence of this sequence.

The single stranded oligonucleotide components are synthesized and annealed in 96 or 384 well plates to generate the double stranded oligonucleotides at a final concentration of 50 pmol/ μ l, 100 μ l total volume per well (Sigma).

2 μ l annealing buffer (NEBuffer 2, 10x concentrated, Biolabs) to a 96 well PCR plate was added to 18 μ l oligos, the plates are spun down briefly and subsequently sealed. The plates are incubated in the PCR machine for 5' at 95°C and slowly cooled. The annealed oligos are diluted 1000x.

Design of pKD122

The knock-down expression vector, pKD122 (figure 3) is based on the pIPspAdapt6 (WO 9964582). The pIPspAdapt6 plasmids contain the 5' part (bp 1-454 and bp 3511-6093 of the adenovirus serotype 5 genome in which the E1 gene is deleted and a promoter is introduced. In contrast, to the plasmid pIPspAdapt6, the siRNA expression vectors of pIPspAdapt lack the CMV promoter and the SV40 polyadenylation site and the larger part of the polylinker, pKD122 further

contains U6 promoter, Sap I recognition sites and the E.Coli lethal gene, ccdB.

Sap I cuts adjacent to its recognition sites (GCTCTTC(N)_{1/4}) creating a 3' overhang (see figure 2). This
5 has the advantage that it cuts any sequence containing the recognition sequence but the recognition sequence will not be present in the final construct since as it will be present on the excised fragment. For these reasons SapI is used for the construction of expression plasmids. The ccdB, is included in
10 the fragment to be excised. When the restriction fragment is not correctly excised and the ccdB gene remains in the plasmid, after transfection no E. coli colonies are formed. Only E. coli containing correct expression plasmids with the two unique SapI overhangs and without the ccdB gene will form
15 colonies.

pIPspAdapt6 was grown in the methylase negative E. coli strain DM1 to prevent methylation of the second Xba-site. The DNA was isolated and digested with Xba I, thereby excising a 142 bp fragment containing the poly A signal. The
20 religated vector is called pIPspAdapt6-deltaPolyA. The polylinker was removed from pIPspAdapt6-deltaPolyA by digestion with EcoRI and BamHI, blunted with Klenow, religated and digested with AscI to reduce background This religated vector is called pIPspAdapt6-deltaPolyA delta-
25 polylinker.

pIPspAdapt6-deltaPolyA-delta polylinker was digested with AvrII and HindIII to remove the CMV promoter and purified on a 1% agarose TAE gel and isolated using the Qiaquick gel extraction kit (Qiagen).

30 The ccdB gene is cut from pIPspAdapt10ZeoDestA (WO9964582) with BamHI and Sal I. The 676 bp ccdB fragment is purified on a 0.8% agarose TAE gel and isolated using the Qiaquick gel extraction kit (Qiagen).

The genomic human U6 gene (Accession number M14486 (GenBank, NCBI)) is cloned by a PCR based strategy using human genomic DNA. The region to be cloned starts at nucleotide -265 upstream of the transcription start site
5 until nucleotide +198 downstream of the transcription start site. The primers used are:

5'-GcacgTTCTAGAAGGTCGGGCAGGAAGAGGGCCT-3' (SEQ ID NO: 11889)

5'-ccgtgcAAGCTTTGGTAAACCGTGCACCGGCGTA-3' (SEQ ID NO: 10 11890)

The PCR product is cloned into the Xba I and Hind III sites of pIPspAdapt6-deltaPolyA, the resulting vector is hU6(+1)pIPspAdapt6-dpA.

Two U6 Sap I PCR fragments (a left, L, and a right R) containing the U6 promoter sequences together with the SapI recognition sequences are made with the following primers:

5'-CGACCATGCGCGGATCCGCTCTTCTGGTGTTCGTCCTT-3' (SEQ ID NO: 11891)

5'-CGGATCCGCGCATGGTCGACGCTCTTCATTACATCAGGTTGTTT-3' (SEQ ID NO: 20 11892).

(SEQ ID NO: 11891) with (SEQ ID NO: 11889) gives the L fragment and (SEQ ID NO: 11892) with (SEQ ID NO: 11890) gives the R fragment with the hU6(+1) pIPspAdapt6-dpA delta polylinker as template.

25 The PCR fragments are purified on a 1% agarose TAE gel and isolated using the Qiaquick gel extraction kit (Qiagen).

The R-fragment is digested with XbaI and BamHI, and the L-fragment is digested with SalI and HindIII. pIPspAdapt6-deltaPolyA-delta polylinker was digested with AvrII and
30 HindIII. The digested R- and L-fragments together with the digested pIPspAdapt6-deltaPolyA delta polylinker and the ccdB fragment are ligated with T4 in ligase buffer (about 30 ng of

each fragment in the ligation) and transformed in DB3.1 cells (wherein the ccdB is not toxic, Invitrogen)

A colony PCR is performed to check sequences with primers SEQ ID NO: 11889 and 11890 this should generate a
5 1000 bp fragment. Positive clones are digested with HincII and BglII, the correct clones give fragments of 3800, 1400, 538, 402 and 134 bp in size. Clones that give these fragments are sequenced. The resulting vector is pKD122 (figure 3)

Cloning of the oligos

10 The knockdown vector pKD122 (figure 3) is digested by Sap I and gel purified. Digestion mix: 30 µl Neb 4, 10 µl Sap I in 300 µl total volume for 9 µg of knock-down vector, and incubated at 37 °C over night. Gel: 1 % agarose in 1x TAE, 2 µl 10x loading buffer to 5 µl of digestion mix; the digested
15 vector is isolated from gel with QIAquick gel extraction kit (Qiagen).

Ligation of the annealed oligos in the knock-down vector: 0.5 µl digested knock-down vector (40 ng/µl), 1 µl T4 DNA ligase buffer (10x concentrated, Biolabs) 0.5 µl T4 DNA
20 ligase (Biolabs) and 7 µl milliQ are added per well. Added to this is 1 µl of the diluted annealed oligos. The plates are incubated over night at RT.

Transformation:

5 µl of each ligation mix is put into a new PCR plate
25 and put on ice 25 µl competent DH5α-cells (Subcloning efficiency, Invitrogen) is added and incubated on ice for 30 minutes. The bacteria are heat shocked for 40 seconds at 37°C and put on ice for 2 minutes. 170 µl RT SOC-medium (Invitrogen) is added to each well. The bacteria are
30 recovered by shaking for 1 hour at 37°C and 100-150 rpm. Cells are spun down at 1700 g for 1 min. 100 µl supernatant is taken and the bacteria are resuspended in the remainder of

the supernatant (100 µl). 50 µl of the cell suspension (50%) is plated out in 1 well of a 6-wells plate (filled with 3 ml LB agar+100 µg/ml ampicillin /well). The plates are incubated overnight at 37°C.

5 **Colony picking:**

3 colonies of each construct are picked and inoculated as agar-stab (LB agar with 100 µg/ml ampicillin) and liquid culture (LB medium with 100 µg/ml ampicillin). The clones in the agar-stab are sequenced. Clones with the correct sequence are transferred to a new 96 well plate. The knock-down vector with the annealed oligos is digested with PI-PspI.

PI-PspI digestion mix (1x) contains 0.5 µg annealed knock-down vector 2.5 µl PI-PspI enzyme, (1U/µl Biolabs), 2.5 µl PI-PspI NEBuffer, (10x concentrated Biolabs), 0.25 µl BSA (100x concentrated Biolabs) in a total volume of 25 µl (end concentration = ± 20 ng/µl). The mixture is incubated over night at 65°C in a humified box. The digestion is checked on gel: 10 µl PI-PspI digestion mix, added to this is 2 µl loading buffer (10x concentrated), and put on gel (1% agarose in 1x TAE buffer + etidium bromide).

20 **Transfection**

Co-transfect each clone to PER.C6/E2A cells together with the cosmid pWE/Ad.AflIII-rITRAE2A (WO9964582. US 6,340,595. Score CPE 14 days after transfection. After the final scoring store the plates at -80°C until further propagation of the viruses.

25 **Virus propagation**

The final virus propagation step is aimed at obtaining a higher percentage of wells showing CPE and more homogenous virus titers. Viruses are propagated according to the following procedure. The transfection plates stored at -80°C

are thawed at room temperature for about 1 hour. By means of a 96 channel Hydra dispenser (Robbins), 20 μ l of the supernatant is transferred onto PER.C6/E2A cells seeded in 96 well plates at a density of 2.25×10^4 cells/well in 180 μ l of DMEM + 10% FBS. Cells are incubated at 34°C, 10% CO₂ for approximately 10 days and the number of wells showing CPE is scored. In general, the number of wells showing CPE is increased after propagation. The plates are then stored at -80°C.

10 In addition, modifications to the viral coat proteins can be introduced to obtain a different or improved tropism (EP 1191105).

The individual knockdown adenoviruses can be used as arrays but also can be pooled to various degrees i.e. sets of pools or one large pool.

Example 3: Knock-down effects with the selected knock-down target sequences:

Knock-down of GNAS:

20 GNAS encodes the G α subunit of G_s, a heterotrimeric G-protein (α , β , γ). G proteins interact with 7 transmembrane receptors, the so-called G-protein coupled receptor (GPCR). Binding of a ligand to the GPCR induces a conformational change of the receptor that results in activation of the G-protein. The activated G protein will dissociate into its α subunit and the $\beta\gamma$ subunit. In the case of G α s the α subunit will interact with adenylyate cyclase and in turn activate this enzyme. Adenylyate cyclase converts ATP into cAMP. Activation of GPCRs that are coupled to G_s will therefore upon activation elevate the cellular cAMP levels. Variation in cAMP levels in the cell can be measured by cAMP responsive elements (CRE). cAMP responsive element activates

transcription when bound by activated cAMP responsive element binding protein (CREBP). CREBP is activated by Protein Kinase A (PKA) that in turn is activated by cAMP. When the CRE is coupled to a luciferase gene, variations in cAMP levels in a cell can be visualized. An increase in cellular cAMP will result in an increase in the amount of luciferase and a decrease of cAMP results in a decrease of luciferase content.

The following assay is designed to measure the knock-down of GNAS at a functional level.

U2OS cells are infected with an adenoviral construct encoding a GPCR ($\beta 2$ adrenergic receptor) that couples to G_s (MOI 500) together with an adenoviral reporter construct carrying CRE elements upstream of a luciferase gene (MOI 750). The infected cells are co-infected with either an adenoviral construct encoding a RNA molecule (SEQ ID NO:1342) targeted against GNAS or an empty virus (MOI 1500).

Two days post infection, the $\beta 2$ adrenergic receptor is activated with isoproterenol, which is an agonist for the $\beta 2$ adrenergic receptor. Six hours later the luciferase activity is measured. After activation of the receptor, successful knock-down of GNAS results in a much lower cAMP levels compared with cAMP levels observed in the control cells transfected with empty virus. To ensure that the decreased cAMP levels are due to the knock-down of GNAS and not due to a non-specific effect on cAMP levels, forskolin is added to the triply infected cells. Forskolin increases intracellular cAMP levels by direct activation of adenylate cyclase and therefore is independent of G-protein (GNAS).

Figure 4 shows the results of the functional knock-down measurements of GNAS. The luciferase activity in cells infected with the adenoviral GNAS knock-down construct are much lower than the luciferase levels of the cells that are

not infected with the GNAS knock-down construct. The forskolin results show that the cAMP reduction is due to the knock-down of GNAS rather than a general down-stream effect on the signal transduction route leading to the activation of CREBP.

Knock-down of SYK

Syk (NM_003177) is a tyrosine kinase that is essential for mast cell degranulation mediated by IgE. In a paper by Moriya et al (1997), it is shown that a specific inhibitor abolished IgE-mediated mast cell degranulation. If a knock-down construct against Syk is successful then mast cells are unable to degranulate.

Mast cell degranulation can be determined by measuring the release of β -hexosaminidase. Total hexosaminidase is measured from the hydrolysis of the synthetic substrate 4-methylumbelliferyl N-acetyl- β -D-glucosaminide by hexosaminidase, which releases fluorescent 4-methylumbelliferone.

Experimental:

Culture of human Mast cells (hMC's)

Human mast cells are derived from cord blood mononuclear cells (#2C-150A; 10^8 cells/vial, Cambrex). Cells are cultured in RPMI 1640 (#52400-025, Gibco BRL) supplemented with 10% FBSHI (Gibco BRL), 1% MEM non essential amino acids (#11140-35, Gibco BRL), 1% L-Glutamin (200mM, 100x) (#25030-024, Gibco BRL), 1% Pen/Strep (10000U: μ g/ml) (#15140-122, Gibco BRL), 0.1% Gentamicin (10mg/ml) (#15710-049, Gibco BRL) and 0.1% 2-Mercaptoethanol (#31350-010, Gibco BRL) in the presence of 100 ng/ml SCF (#300-07, Peprotech), 50 ng/ml hIL-6 (#200-06, Peprotech) and 10 ng/ml hIL-10 (#200-10, Peprotech). The cell suspensions are seeded at a density of 10^6 cells/ml and the entire volume of cytokine supplemented medium is replaced on a weekly basis. The adherent fraction

of cells is discarded weekly by the transfer of the non-adherent cells to fresh culture flasks. Cells are cultured for 9-11 weeks before use as described by Ochi et al (1999)

Two weeks prior to the experiment cells are cultured
5 with addition of 10 ng/ml hIL-4 and 5 µg/ml human myeloma IgE (Biodesign) to enhance FcγRI expression as described by Ochi et al (2000). Cells are checked by toluidine blue staining and Flow cytometric analysis. Toluidine blue staining is performed on a weekly basis starting 6th week of culture.
10 Cells are stained for C-kit and FcεRI expression and analyzed by FACS.

Degranulation assay

Cells are seeded at a density of 1E+05 cells/well in a 96 well V-bottom plate (Greiner) and transduced with an
15 Adenovirus at a MOI of 1000. Three days past infection the culture medium containing virus is removed from the cells and replaced by 200 µl of 1500 ng/ml anti-IgE (#501, DAKO) in culture medium. After one-hour of stimulation, at 37°C, 100 µl supernatant is transferred to a (black) flat-bottom 96
20 well plate (Greiner). The remaining supernatant is discarded and the cells are lysed (to determine β-hexosaminidase present in cells) using 200 µl 0.2% Triton X-100 (#B32844, Calbiochem) in culture medium. After cell lysis 100 µl supernatant is transferred to a (black) flat-bottom 96 well
25 plate. For fluorimetric (Excitation at 360nm/Emission at 450nm) determination, 50 µl of 4mM 4-methylumbelliferyl N-acetyl-β-D-glucosaminide (4-MUG) (Sigma) is added to the supernatant on both plates. The plates are incubated for 1 hour at 37°C after which 100 µl of 0.4 M glycine buffer (pH
30 10.7) is added to stop the reaction. The hydrolysis of 4-MUG is measured with excitation at 360nm and emission measured at: 450nm. The degranulation is represented as the percentage of

β -hexosaminidase released from the total amount of β hexosaminidase present in the cells before stimulation.

Experimental setup

SykT402 a dominant negative form of Syk, which is a key
5 player in the degranulation pathway, as shown by Moriya et al
(1997). SykT402 is a truncated form of WT Syk at amino acid
402. Cells with the SykT402 mutant are unable to degranulate.

In this experiment, cells are transduced, at a MOI of
1000.

10 Cells are transduced either with the SykT402 mutant, a
Syk knock-down virus construct (SEQ ID NO: 1338 and see
example 2) or empty virus as a negative control. Knock-down
virus directed against GL2 is taken along to check for
specific knock-down. Three days past transduction, cells are
15 stimulated with 1500 ng/ml anti-IgE as described in
'degranulation assay'. Without stimulation the mast cells
have already a β -hexosaminidase release of ~7%, this
background is subtracted from all values measured. When
comparing anti-IgE only (no viral infection) to empty (virus)
20 one can see that viral infection has little to no effect on
degranulation.

Results:

The results are summarized in figure 5. As expected, the
truncated form of Syk (SykT402) shows a low level of β -
25 hexosaminidase release (about 20% of normal levels). The
control virus for knockdown, GL2, shows a comparable β -
hexosaminidase release level as the empty virus,
demonstrating that there is no non-specific knock-down
effect. The knockdown virus g-SYK (63) displays a level of β -
30 hexosaminidase release similar to the dominant negative
mutant of Syk (SykT402), demonstrating the effectiveness of
our knock-down constructs.

Knock-down of IKK- β

The functional effect of an adenoviral knock-down was studied using the well-known NF κ B pathway by measuring the phosphorylation of I κ B α , one of the inhibitory proteins which
5 normally associate with and hold NF- κ B inactive in the cytoplasm. Activation of the IKK kinase complex by TNF α stimulation results in phosphorylation of I κ B β , followed by its ubiquitination and degradation by the 26S proteasome. Phosphorylation and degradation of I κ B α releases NF- κ B,
10 allowing it to translocate into the nucleus to activate target genes.

In the following experiment, U2OS cells were infected with IKK β knock-down virus (SEQ ID NO:1343), GL2 control knock-down virus, or the cells were not infected at all.
15 Figure 6 clearly shows that the adenoviral siRNA expression gives a specific reduction of the IKK β protein, but leaves β -tubulin unaffected (6 days post infection). At 6 days post-infection, the cells were stimulated with 50 ng/ml TNF α for 0 minute, 10 minutes, or 30 minutes. The phosphorylation state
20 of I κ B α was determined by Bio-Plex phosphoprotein analysis using phospho-specific I κ B α antibodies. U2OS cells stimulated with TNF α show strong induction of phosphorylated I κ B α peaking around 10 minutes and fading by 30 minutes. The peak level of induction of phosphorylated I κ B α was reduced in cells that
25 were infected with adenoviral knock-down constructs directed against IKK β . Altogether these data demonstrate that our adenoviral knock-down system results in specific reduction of the target at the levels of mRNA, protein, and function of the protein.

30 Phosphoprotein assay for I κ B α

The phosphorylation state of I κ B α was determined by using the Bio-Plex phospho-I κ B α kit (Bio-Rad), an assay that

is similar to a capture sandwich ELISA but utilizes antibodies covalently coupled to dyed polystyrene beads. U2OS cells were infected with adenoviral knock-downs as indicated. Six days post infection, cell lysates were
5 prepared after either treatment with 50 ng/mL of TNF α (Sigma) for 10 or 30 minutes, or non-treated (0 minute). The Bio-Plex phospho-I κ B α assays were performed according to the manufacturer's instruction. Briefly, equal amount of cellular proteins were incubated overnight with the dyed beads
10 conjugated with the I κ B α capture antibodies. After a series of washes to remove unbound protein, a biotinylated detection antibody specific for Phospho-I κ B α (Ser32) was added to the reaction. Streptavidin-phycoerythrin (PE) was added to detect the formation of a sandwich of antibodies around the
15 phosphorylated I κ B α . The Luminex IS100 system was used to read each individual bead in the reaction mixture (Luminex, Austin).

Knock-down of MMP2

In another experiment, the constitutively expressed
20 matrix metalloproteinase *MMP2* (gelatinase A) was targeted for knock-down with 2 different adenoviral constructs (Ad-siRNA-MMP2-1 (SEQ ID NO: 633 or Ad-siRNA-MMP2-3 SEQ ID NO: 635) in primary human synoviocytes. Supernatants of primary
synoviocytes were analysed 6 days after infection at MOI 7500
25 with Ad-siRNA-MMP2-1 and Ad-siRNA-MMP2-3, and compared to Ad-siRNA-control, a virus containing a target sequence against *M6PR*. Both constructs resulted in very efficient reduction (>95% knock-down) of *MMP2* mRNA determined by real time PCR analysis (figure 7), as well as the *MMP2* protein activity as
30 determined by its gelatin-degrading activity using zymogram gels (figure 7). Adenoviral overexpression of the *MMP2* protein (Ad-MMP2) demonstrates the identity of the band

visualised in this assay. Altogether, this clearly shows that
adenoviral vectors can efficiently introduce and express
siRNAs to levels that result in a specific functional knock-
down of the targeted transcript and the encoded protein in
5 primary human cells.

References:

- Curriel DT, Wagner E, Cotten M, Birnstiel ML, Agarwal S, Li CM, Loechel S, Hu PC. High-efficiency gene transfer mediated by adenovirus coupled to DNA-polylysine complexes. Hum Gene Ther. 1992, 3(2):147-54.
- Elbashir SM, Harborth J, Weber K, Tuschl T. Analysis of gene function in somatic mammalian cells using small interfering RNAs. Methods. 2002, 26(2):199-213.
- Jarvis and Ford (2002) Ambion TechNotes 8(5), p 852
- Elbashir SM, Harborth J, Lendeckel W, Yalcin A, Weber K, Tuschl T. Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. Nature. 2001 411(6836):494-8.
- Katsuhiko Moriya, Juan Rivera, Sandra Odom, Yoshinori Sakuma, Kenzo Muramoto, Tatsuya Yoshiuchi, Mitsuaki Miyamoto, and Koji Yamada. ER-27319, an acridone-related compound, inhibits release of antigen-induced allergic mediators from mast cells by selective inhibition of Fcε receptor I-mediated activation of Syk. Proc Natl Acad Sci U S A (1997) 94: 12539-12544
- Manning, G, Whyte, D. B., Martinez, R., Hunter, T., Sudarsanam, S. The Protein Kinase Complement of the Human Genome Science 2002: 1912-1934.
- Moriya K, Rivera J, Odom S, Sakuma Y, Muramoto K, Yoshiuchi T, Miyamoto M, Yamada K. ER-27319, an acridone-related compound, inhibits release of antigen-induced allergic mediators from mast cells by selective inhibition of fcepsilon receptor I-mediated activation of Syk. Proc Natl Acad Sci U S A. 1997, 94(23):12539-44.
- Ochi H, De Jesus NH, Hsieh FH, Austen KF, Boyce JA. IL-4 and -5 prime human mast cells for different profiles of IgE-

NM_022450	FLJ22357	8316	ACTGGCAGCGCAAGAGCATCC
NM_022450	FLJ22357	8317	ACGTACGTGCAGCAGGAGAAC
NM_022450	FLJ22357	8318	ACCCTGGCCAGTGCCATCTTC
NM_004181	UHL1	8319	ACGCAGACCATTGGGAATTCC
NM_004181	UHL1	8320	ACTGAGGCCATACAGGCAGCC
NM_004181	UHL1	8321	ACGGTCTGCAGAGAATTCAAC
NM_003368	USP1	8322	ACCACACTGAGGGAACTCAAC
NM_003368	USP1	8323	ACGTCTGCAACTAAGCAACCC
NM_003368	USP1	8324	ACTGAAGACCCCTGAGATGGGC
NM_005747	ELA3A	8325	ACCCTTGCTGTGAAGGAGGGC
NM_001948	DUT	8326	ACTACCACCTATGGAGAAAGC
NM_014122	PRO0245	8326	ACTACCACCTATGGAGAAAGC
NM_012464	TLL1	8327	ACTTGACCTTACGCAGAACCC
NM_012464	TLL1	8328	ACAGAGGAAGATGAGTGTGCC
NM_012464	TLL1	8329	ACCTACCCAGGACAGGTTGAC
NM_015670	SENP3	8330	ACAGCACCTCGCTGACATTCC
NM_015670	SENP3	8331	ACGTCTCCTCTGGACCCTGAC
NM_015670	SENP3	8332	ACGACCGACTGGATTTCACC
NM_017435	SLC21A14	8333	ACGCATTGGCAGAAGGCTATC
NM_017435	SLC21A14	8334	ACCTGTCTCCTACCAAGGAAC
NM_017435	SLC21A14	8335	ACGAGTTCTTGCAGGAATCCC
NM_018561	DKFZP586D2223	8336	ACCCTGGGCAACACCTGCTAC
NM_018561	DKFZP586D2223	8337	ACTTCACAGAGACAGAGGCCC
NM_003879	CFLAR	8338	ACCCCTCACCTTGTTTCGGAC
NM_003879	CFLAR	8339	ACTTTGCCTGTATGCCCGAGC
NM_003879	CFLAR	8340	ACCTATGTGGTGTGAGAGGGC
NM_032823	FLJ14675	8341	ACCTATGCCAATGCCAGCCTC
NM_032823	FLJ14675	8342	ACCTGGCCACACAAGTGACTC
NM_032823	FLJ14675	8343	ACTGCTACTGGAGAACATTCC
NM_024539	FLJ23516	8344	ACCATGGCCCTTGGGTGAATC
NM_024539	FLJ23516	8345	ACGGAGACAAGGAAATTGGCC
NM_144575	CAPN13	8346	ACGATCCTGATGGTCCAAAGC
NM_144575	CAPN13	8347	ACGCTGCTCGGATCCTATTCC
NM_144575	CAPN13	8348	ACGTGATTCTGGCTGGCTCAC
NM_144691	MGC20576	8349	ACCAAGTTCGATGAGGACACC
NM_144691	MGC20576	8350	ACCAACCAGCTGACCCAGACC
NM_004793	PRSS15	8351	ACGCAGACCCACCGTAAGTAC
NM_004793	PRSS15	8352	ACGAAACGCATCCTGGAGTTC
NM_004793	PRSS15	8353	ACGGAGAAGACCATTGCGGCC
NM_005865	PRSS16	8354	ACTGACCAACATTGGGTTGGC
NM_005865	PRSS16	8355	ACTGACGTGGTATCCCGAAGC
NM_005865	PRSS16	8356	ACCACAGGCTTTAGGATCCTC
NM_144981	FLJ25059	8357	ACTCCTCACCAGTAGTCCATC
NM_144981	FLJ25059	8358	ACTTCTACAGATTCCAGGTGC
NM_017414	USP18	8359	ACCATGAAGAGAGAGCAGCCC
NM_017414	USP18	8360	ACCCTGATTAAGGACCAGATC

Claims

1. Polynucleotide comprising an RNA sequence comprising a first stretch of 21 consecutive nucleotides and a second
5 stretch of 21 consecutive nucleotides, complementary to the first stretch of 21 consecutive nucleotides, wherein in first stretch of 21 consecutive nucleotides:
- a. in the 5' -> 3' direction, the first nucleotide is an A-nucleotide , the second nucleotide is a C-nucleotide
10 and the last nucleotide is a C-nucleotide; and
 - b. no stretches of four or more consecutive identical nucleotides are present;
 - c. the total number of G- and C-nucleotides is between 33-71% of the total number of nucleotides
 - 15 d. in the 5' -> 3' direction, consecutive nucleotides 3-21 are homologous to a RNA-molecule.
2. Polynucleotide according to claim 1, wherein the RNA-molecule is a human RNA molecule.
3. Polynucleotide according to claims 1-2, wherein in
20 the first stretch of 21 consecutive nucleotides no stretches of three or more consecutive A- nucleotides are present
4. Polynucleotide according to claims 1-3 wherein in the first stretch of 21 consecutive nucleotides no stretches of three or more consecutive U-nucleotides are present
- 25 5. Polynucleotide according to claims 1-4, wherein, in the 5' -> 3' direction, consecutive nucleotides 3-21 of the first stretch of 21 consecutive nucleotides are unique.
6. Polynucleotide according to claims 1-5, wherein, in the 5' -> 3' direction, consecutive nucleotides 3-21 of the
30 first stretch of 21 consecutive nucleotides are homologous to a sequence positioned at least 75 nucleotides downstream of the translation initiation site of the transcribed RNA molecule encoding a polypeptide.

7. Polynucleotide according to claims 1-6, wherein, in the 5' -> 3' direction, consecutive nucleotides 3-21 of the first stretch of 21 consecutive nucleotides are homologous to a sequence positioned at least upstream of the translation
5 termination site of the transcribed RNA molecule encoding a polypeptide.

8. Polynucleotide according to claims 1-7 wherein the RNA sequence also comprises a linker sequence linking the first stretch of 21 consecutive nucleotides with the second
10 stretch of 21 consecutive nucleotides.

9. Polynucleotide according to claim 8, wherein the linker sequence is 4-30 nucleotides long, preferably 5-15 nucleotides long and most preferably 8 nucleotides long

10. Polynucleotide according to claims 8-9, wherein the
15 linker sequence is SEQ ID NO: 1

11. Polynucleotide according to claims 1-10, wherein the first stretch of 21 consecutive nucleotides is selected from a group consisting of SEQ ID NO: 2-11888

12. Polynucleotide according to claims 1-11, wherein the
20 first stretch of 21 consecutive nucleotides is selected from a group consisting of SEQ ID NO: 1342, 1338, 1343, 633, 635

13. Vector capable of transfecting a host cell and comprising a sequence encoding the polynucleotide according to claims 1-12 and a promoter sequence operatively linked to
25 the sequence encoding the polynucleotide.

14. Vector according to claim 13, wherein the promoter is a microRNA promoter, preferably a let-7 promoter.

15. Vector according to claim 14, wherein the promoter is a promoter recognized by RNA Polymerase III, preferably U6
30 small nuclear RNA.

16. The vector according to claims 13-15, wherein the vector is an adenoviral vector, preferably the adenoviral vector is replication defective.

17. Library of polynucleotide sequences according to claims 1-12.

18. Library of vectors according to claims 13-16.

19. Library according to claim 18 wherein the vectors
5 are viral vectors preferably selected from a group consisting of AAV, Lentivirus or Retrovirus.

20. Library according to claim 18 wherein the vectors are adenoviral vectors, preferably the adenoviral vectors are replication defective.

10 21. Method of making a vector according to claim 16 comprising the steps:

a. synthesizing a forward primer with the following sequence in the 5' -> 3' direction:

i. the nucleotides ACC

15 ii. a DNA sequence corresponding to the antisense sequence of a sequence selected from the group consisting of SEQ ID NO: 2-11888

iii. the nucleotides TTGCTATA

20 iv. the antisense sequence corresponding to the DNA sequence from step a-ii

v. the nucleotides TTT

b. synthesizing a reverse primer with the following sequence in the 5' -> 3' direction

i. the nucleotides TAAAAA

25 ii. a DNA sequence corresponding the antisense sequence of a sequence selected from the group consisting of SEQ ID NO: 2-11888

iii. the nucleotides TATAGCAA

30 iv. the antisense sequence corresponding to the DNA sequence from step b-ii

c. annealing the primers from step (a) and (b)

d. exchanging the ccdB sequences in plasmid pKD122 for the annealed primers from step (c)

22. Method of preparing a vector according to claim 15 comprising the steps:

a. synthesizing a forward primer with the following sequence in the 5' -> 3' direction:

- 5 i. the nucleotides ACC
- ii. a DNA sequence corresponding to a sequence selected from the group consisting of SEQ ID NO: 2-11888
- iii. the nucleotides TTGCTATA
- iv the antisense sequence corresponding to the DNA
10 sequence from step a-ii
- v. the nucleotides TTT
- b. synthesizing a reverse primer with the following sequence in the 5' -> 3' direction
- i. the nucleotides TAAAAA
- 15 ii. a DNA sequence corresponding to a sequence selected from the group consisting of SEQ ID NO: 2-11888
- iii. the nucleotides TATAGCAA
- iv. the antisense sequence corresponding to the DNA sequence from step b-ii
- 20 c. annealing the primers from step (a) and (b)
- d. exchanging the ccdB sequences in plasmid pKD122 for the annealed primers from step (c).

23. Method of determining the function of a naturally occurring polynucleotide sequence comprising transfecting a
25 host cell with a vector according to claims 13-16, the vector transcribing a polynucleotide sequence according to claims 1-12 and detecting a change in cellular phenotype.

24. Method of determining the function of a naturally occurring polynucleotide sequence in a high throughput
30 setting,

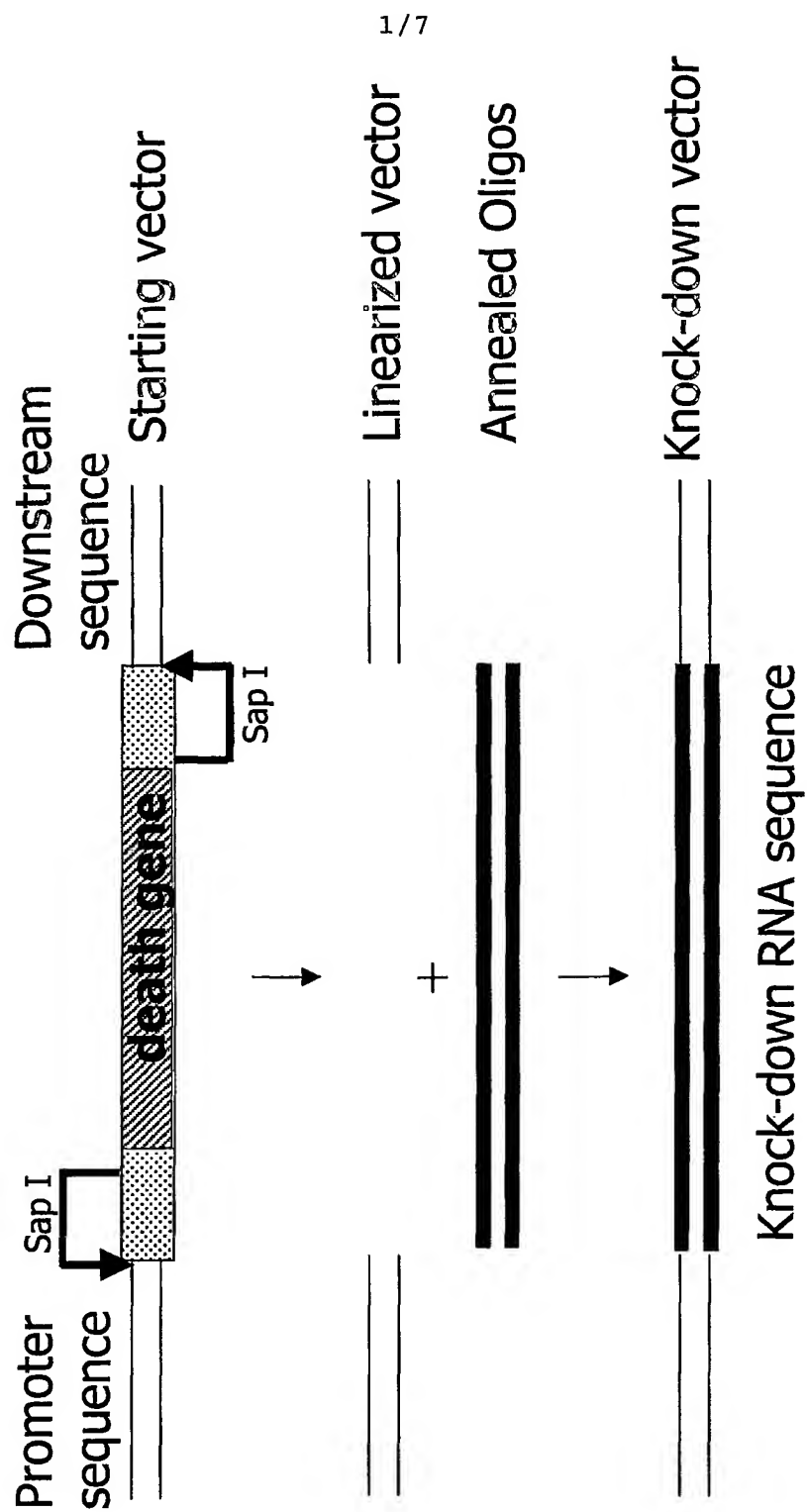
a. providing a library of vectors according to claims
17-20

b. transducing a host cell with the vectors of step (a),

- c. expressing in the host cell the product(s) of the vectors of step (a),
- d. thereby altering a phenotype of the host,
- e. identifying the altered phenotype and,
- 5 f. assigning a function to the naturally occurring polynucleotide sequence (s).

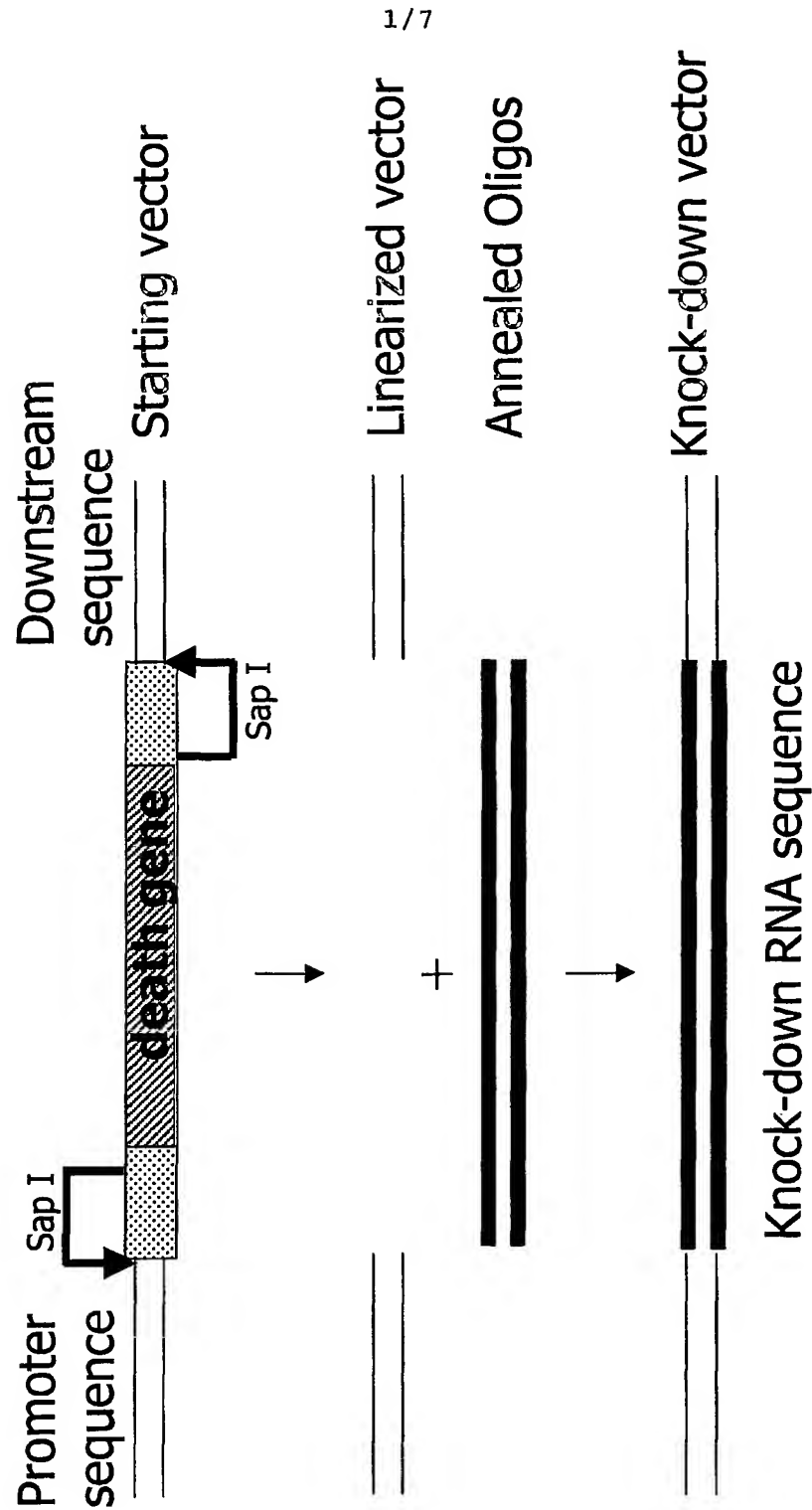
Vector development

Figure 1



Vector development

Figure 1



Adenovirus vector development

Knock-down cloning vector

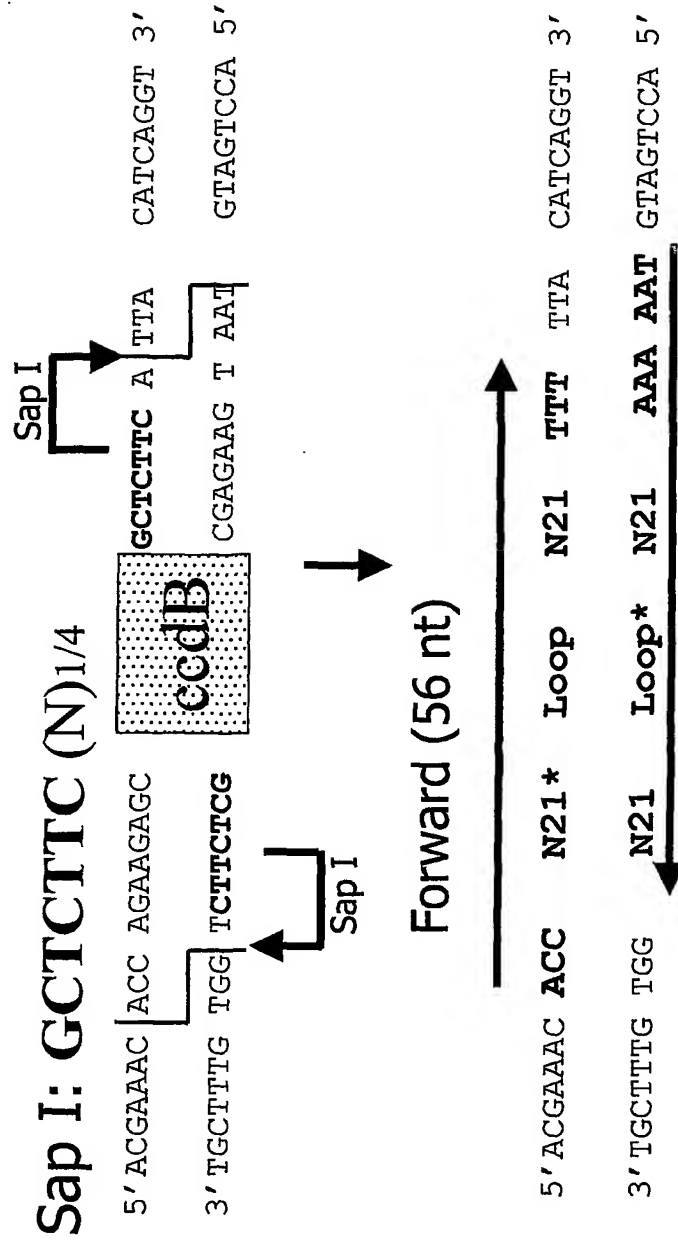


Figure 2

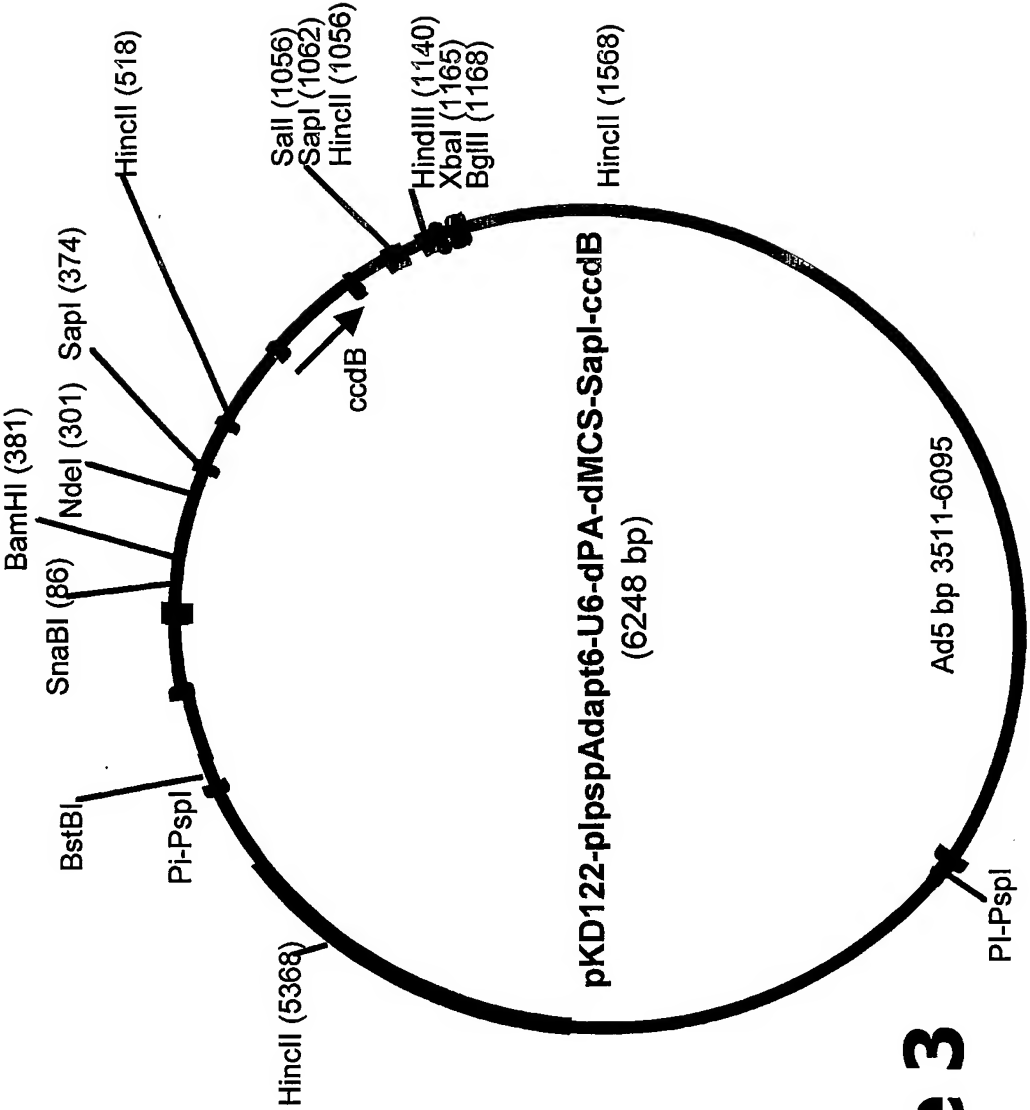


Figure 3

Knock-down of GNAS

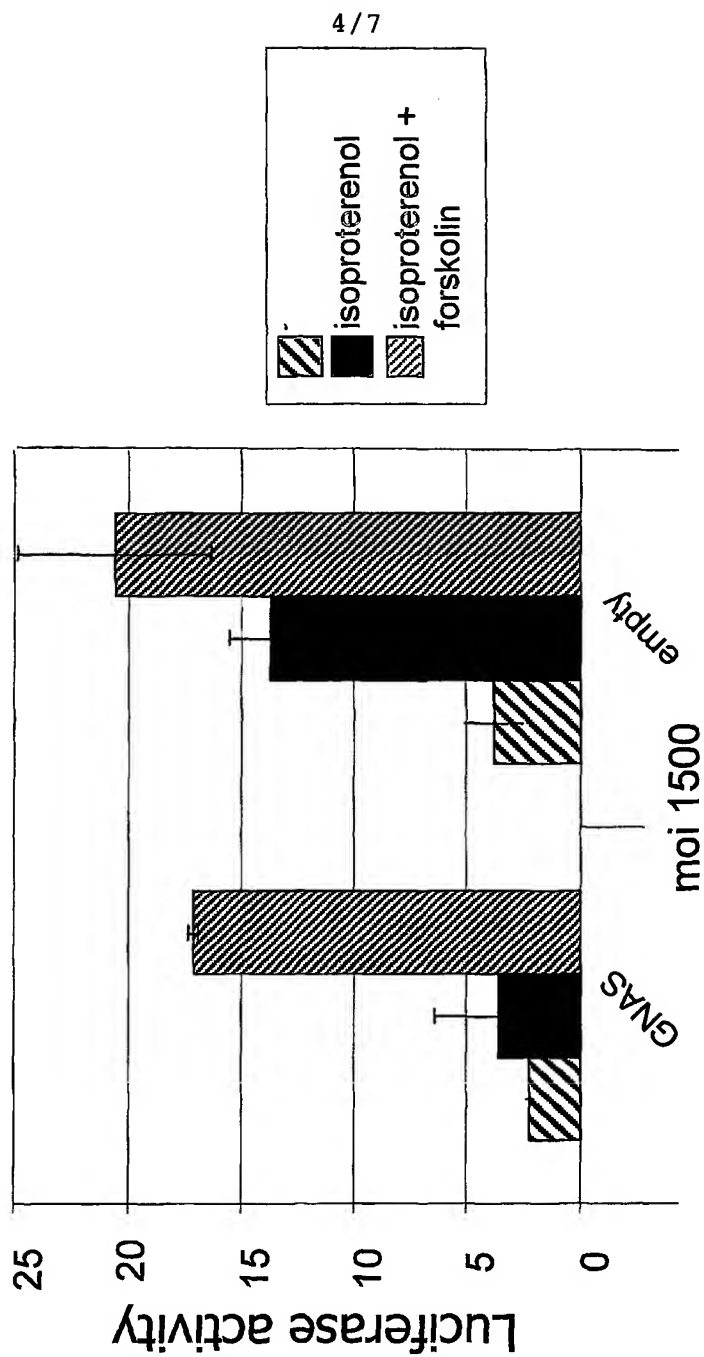


Figure 4

Knock-down of SYK Released Beta-hexosaminidase

Released beta-hexosaminidase Human Mast Cells
triplo, (background subtracted)

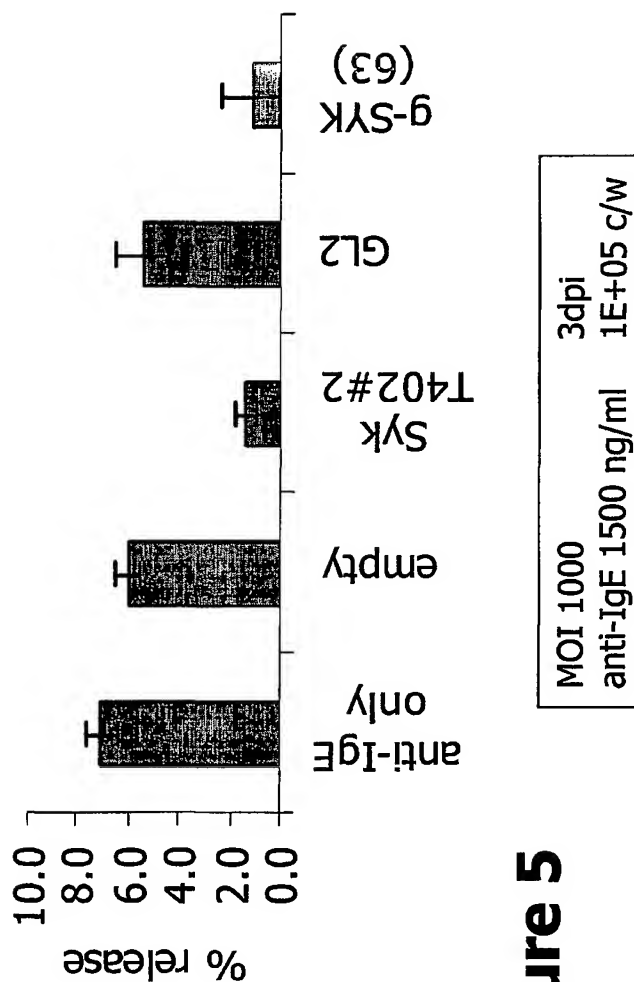
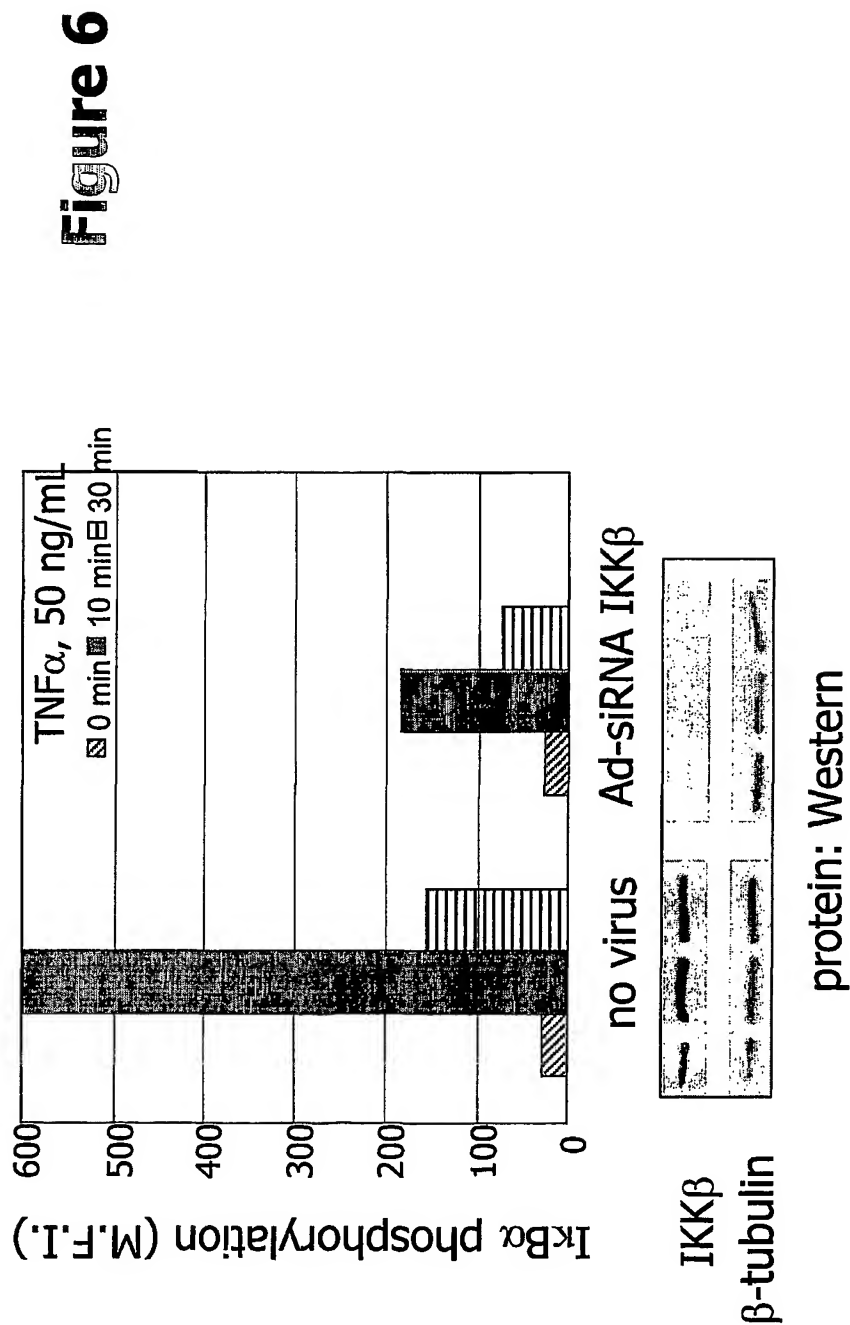


Figure 5

TNF α induced phospho-I κ B α **Effect of Ad-IKK β knock-down on phospho-I κ B α**



MMP-2 knock-down

Reduction of MMP-2 mRNA and activity levels

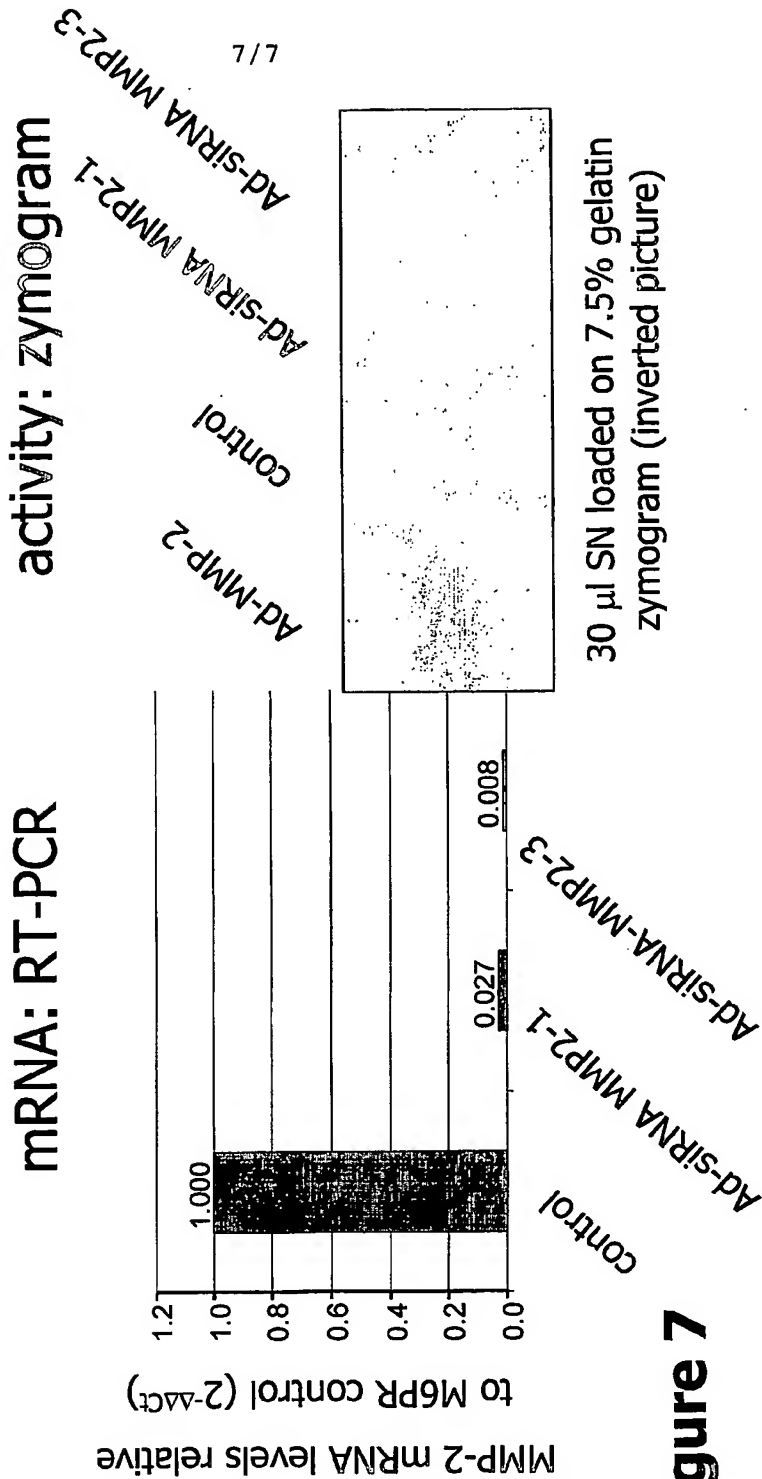


Figure 7

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 03/04362

A. CLASSIFICATION OF SUBJECT MATTER

IPC 7 C12N15/11 A01K67/027 C12N15/861 C12Q1/68 C12N15/10
C12N5/10

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 C12N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the International search (name of data base and, where practical, search terms used)

CHEM ABS Data, EPO-Internal, WPI Data, PAJ, BIOSIS, MEDLINE

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 03 020931 A (GALAPAGOS GENOMICS N V ;ARTS GERT-JAN (NL); PIEST IVO (NL); LANGEM) 13 March 2003 (2003-03-13) cited in the application example 6	1-10,13, 15,16
Y	example 6	11,14, 17-24
Y	figures 25A,25B,26-28; examples 3,4,9,11,12	11,14, 17-24
X	WO 03 012052 A (ALTON ERIC WFW ;GRIESENBACH UTA (GB); IMP COLLEGE INNOVATIONS LTD) 13 February 2003 (2003-02-13) SEQ ID NO's: 264 & 265 example 6; table 8	1-7

--/--

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents:

- *A* document defining the general state of the art which is not considered to be of particular relevance
- *E* earlier document but published on or after the international filing date
- *L* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- *O* document referring to an oral disclosure, use, exhibition or other means
- *P* document published prior to the international filing date but later than the priority date claimed

- *T* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- *X* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- *Y* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- *G* document member of the same patent family

Date of the actual completion of the international search

12 August 2003

Date of mailing of the international search report

04.12.03

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,
Fax: (+31-70) 340-3016

Authorized officer

Steffen, P